Sciences Po
Master of Public Affairs 2009-2010

Syllabus, Required Course, 1st semester (2.5 credits)

Statistical and Data Analysis for Policymakers - Level 1A

Bruno Cautrès

Instructor :

Dr. Bruno Cautrès
CEVIPOF- Centre de recherches politiques de Sciences Po
98 rue de l'Université
75007 – PARIS
bruno.cautres@sciences-po.fr

TA :

Hyungsoo Woo
hyungsoo.woo@sciences-po.org

Course Description:

This Level 1 course is designed to introduce students to statistics and data analysis as a means for analyzing data in the fields of public policy, politics and political economy analysis. The objective is giving a foundation in statistical methods mainly used to describe data.

The course covers topics addressed in most introductory statistics classes, i.e., simple univariate and bivariate statistics for discrete and continuous variables, ending with correlation analysis and possibly (but not compulsory for this introductory course) ordinary regression. It is an applied course with lectures, readings and practical sessions. It emphasizes practice (e.g., choice of appropriate statistical procedures, diagnostics, interpretations) over theory (mathematical derivations and proofs). Mathematical formulation will be kept at the necessary minimum level, the objective is not learning complex formulas but learning how to use it in practical situations.

Graphical display techniques (for instance box-plots, stem and leaves) will also be addressed. Distribution functions and the logic of statistical hypothesis testing will also be covered since they are key elements of the decision-making process.

An important goal of the course is to make students aware of the interaction between statistical validation criteria and public policy/political decision making process.

The course combines lectures and classes in an interactive format.

Assessment:

The course is evaluated on the following basis :

Take-home assessment tasks : you will have 3 assignments, with 2 weeks to complete each of the assignments. The mean of the three assignments will account for 50 % of the final grade.

A final exam that will be organized at the end of the course; this exam will cover questions/answers type of validation and a small practical example. The final exam will account for 50% of your final grade.

Monitoring of the course and of the students' work:

A short meeting between the student delegate, instructor and/or TA and the MPA administration will be organized every two courses, to ensure the quality of the interaction between the course and the students.

Textbook and software:

There are several textbooks that may be of use for introductory methods. Three excellent and instructive books :

- P. Newbold, W.L. Carlson and B.M. Thorne. *Statistics for Business and Economics* (Sixth Edition), Prenctice Hall, 2003.

- Thomas H. Wonnacott and Ronald J. Wonnacott. *Introductory Statistics for Business and Economics* (Fourth Edition). New York, Wiley, 1990.

- Alan Agresti, Barbara Finlay. *Statistical methods for social sciences*. Prentice Hall, 2007, (4^{th} edition).

These books are introductory statistics books that are used in many US and worldwide universities. They cover much more than Level 1 course topics and offer a large scope of statistical procedures. However, there are also many other introductory statistics textbooks that cover the same material. Furthermore, many websites provide excellent introductions to statistics. A list of such sites will be given at the beginning of the course.

In 2009/2010, the Agresti/Finlay textbook will be used as a reference textbook. You may buy it, new or used. In any case, the MPA will provide you copies of the relevant pages. Every session will be based on a PowerPoint presentation, made available to students through an ENTG workgroup.

Statistical software used for this course will be STATA. An excellent introduction to this software is available. A list of possible textbooks about software will also be provided during the first session. Some documents focused on an introduction to the programming STATA command language will be provided by the instructor.

Course outline:

Session 1 : Course Outline. Introduction. Course administration. Key concepts and objectives of introduction to statistical data analysis.

Session 2 : Sampling and measurement.

Session 3 : Statistical description I : using graphical displays.

Session 4 : Statistical description II : summary statistics.

Session 5 : Probability and distributions : the normal distribution, the standard normal distribution, sampling and the central limit theorem.

Session 6: Confidence intervals, estimators and their properties (unbiasedness, consistency and efficiency).

Session 7 : Hypothesis testing I

a) Hypothesis testing logic

b) More on one-sample tests.

Session 8 : Recap and "pop quiz session" : confidence intervals, hypothesis testing

Session 9 : Hypothesis testing II : equality of means and other tests.

Session 10 : Bivariate analysis I : crosstabs and the chi-square test

Session 11 : Bivariate analysis II : correlation analysis

Session 12 : Conclusions